一种 H.265 视频带宽实时预测模型 *

刘江涛，李泽平，杨炳钊
(贵州大学 计算机科学与技术学院，贵阳 550025)

摘 要：针对视频直播过程中带宽抖动导致视频帧丢失造成用户体验降低问题，提出了一种 H。265 视频带宽实时预测(VBRP)模型。模型基于马尔可夫琏，专注于 H.265 编码的直播视频流，研究了在 GOP(group of pictures)编码模式下 B 帧出现的统计特性，发现 B 帧大小对视频流传输速率的影响较大，应用此特性，可预测出直播流中 B 帧的大小，并根据网络带宽选择性丢弃 B 帧，同时引入步长调整因子 AF 和误差阈值 FT，均衡模型在预测 B 帧时的训练频率和预测帧数。最后基于模型实现 VBRP 预测算法并在直播系统中验证算法的有效性。

关键词：带宽预测；H.265；马尔可夫；流媒体
中图分类号：TP393.0 doi: 10.19734/j.issn.1001-3695.2018.08.0552

Real-time prediction model of H.265 video bandwidth

Liu Jiangtao, Li Zeping, Yang Bingzhao
(*College of Computer Science & Technology Guizhou University*, *Guiyang* 550025, *China*)

Abstract: In order to solve the user experience reduction by bandwidth jitter that caused video frame loss during live video, This paper proposed a real-time prediction model (VBRP) of H. 265 video bandwidth. The model is based on markov chain and focuses on the live video stream encoded by H. 265. It studied the statistical characteristics of frame B in the GOP(Group of pictures) encoding mode, which showed that the B frames size had a great influence on the transmission rate of video stream and used this feature, it can predict the size of the B frame, then according to the network bandwidth selective discarding the B frame, at the same time used the number of frames adjustment factor AF and error threshold FT, to balance the training frequency and the number of predicted frames when predicting B frames. Finally, the prediction algorithm is implemented and validated in the live video system.

Key words: bandwidth prediction; H.265; Markov; media streaming

0 引言

根据 Netflix 的 2017 年度报告[1]全世界有超过 1.17 亿的 Netflix 流媒体会员，每天有超过 1.4 亿小时的电视节目和电影被观看。未来，流媒体流量预计将在移动领域增长 67%，在固定网络领域增长 29%[2]。多媒体内容的爆炸式增长推动了流媒体行业的发展和研究，寻求更高效的流媒体发布方法架构为所有用户提供更好的视频服务成为研究热点。在当前的网络服务中，视频流已经占据了绝大部分有线和无线的网络带宽。 H.265/HEVC(high efficiency video coding)视频编码[3]在同等图像质量下，其压缩率比 H.264 标准提高了一倍，所以 H.265/HEVC 编码应用于流媒体传输将会极大提高网络带宽利用率。然而，网络带宽并不总是稳定不变的，通常带宽是一个时变系统，随时间而动态变化，特别是在高抖动、高时延的移动网络环境下尤为明显。如果在视频直播中遇到网络抖动，将造成关键视频数据丢失，影响视频画质，传统的解决方案是根据网络带宽的减少去降低视频的质量[4]，这样就给用户带来不好的体验并且会使服务质量下降。如果用统计学方法预测视频带宽大小，并根据网络带宽选择性丢弃 B 帧。能够平衡服务质量(quality of service，QoS)与网络带宽负载能力，从而提高用户体验(quality of experience，QoE)。

1 相关工作

Sarkar 等人[5]提出了 Markov-modulated gamma-based 框架,Lanfranchi 等人[6]提出了 MPEG-4 预测模型来提高用户的 QoE 和服务提供商的 QoS，这些模型必须适应动态变化的网络条件，如吞吐量、丢包率和延迟抖动[8]。Kalampogia 等人[9]提出的模拟退火算法来预测视频的 B 帧，然而模拟退火算法和 MPEG-4 预测模型和对于 H.265 编码的视频来说是不够准确的。因为 H.265 视频有许多比均值小很多的 B 帧，而且随机分布在整个视频帧中，不能用公式很好地预测。根据文献[5]分析，视频模型可以分为两类，分别是数据速率模型和帧大小模型。数据速率模型不能识别帧对整个视频的影响程度，如果一些包括 I 帧在内的很小数据丢失，会影响接收视频的质量。在文献[10]中，作者计算了视频流数据包在 LTE 网络中准时传输的概率,根据概率将视频数据包分配给用户，而不区分视频帧不同类型在各个视频中的重要性。然而文献[10]遵循的是 DASH 标准[11]，而 DASH 标准不适合非常低延迟的视频流。文献[12]用目前已知的分布对 B 帧进行建模，找到 B 帧的统计特性，但是对于直播视频流，即将到来的视频帧是未知的，无法进行统计，因此会很大程度上影响 B 帧预测的准确性。

收稿日期：2018-08-03；**修回日期：**2018-09-12 **基金项目：**国家自然科学基金资助项目（61462014）；贵州省优秀科技教育人才省长专项资金资助项目（黔省专合字［2011］34）；贵州省科学技术基金资助项目（黔科合 J 字［2011］2201）；贵州大学引进人才科研项目（贵大人基合字［2010］034）

作者简介：刘江涛（1991-），男，四川达州人，硕士研究生，主要研究方向为计算机网络、流媒体技术(jonta@jonta.cn)；李泽平（1964-），男，贵州贵阳人，教授，硕导，博士，主要研究方向为计算机网络、流媒体技术；杨炳钊（1993-），男，山东烟台人，硕士研究生，主要研究方向为计算机网络与流媒体技术.

针对以上问题，提出一种 H.265 视频带宽实时预测(video bandwidth real-time prediction prediction，VBRP)模型。VBRP 模型基于马尔可夫链，用于动态网络环境下提升视频直播的质量。在此模型基础上实现了视频带宽实时预测 VBRP 算法，应用于 RTMP(real time messaging protocol)[7]流媒体发布预测系统。本文专注的 H.265[3]是继承 H.264 的新一代视频编码标准，与 H.264 一样，在视频编码序列中，主要有三种编码帧: I 帧、P 帧、B 帧组成 GOP，I 帧即 intra-coded picture（帧内编码图像帧），不参考其他图像帧，只利用本帧的信息进行编码；P 帧即 predictive-coded picture（预测编码图像帧），利用之前的 I 帧或 P 帧，采用运动预测的方式进行帧间预测编码；B 帧即 bidirectionally predicted picture（双向预测编码图像帧)，提供最高的压缩比，它既需要之前的图像帧(I 帧或 P 帧)，也需要后来的图像帧(P 帧)，采用运动预测的方式进行帧间双向预测编码。本文研究了直播视频流的统计特性，即 GOP 编码模式下 B 帧对传输速率的影响程度，说明了根据当前带宽负载能力可选择地丢弃预测的 B 帧对提高用户体验的意义。

2　H.265 编码的直播视频流特征

H.265 编码的视频流中不但 B 帧的数量占了每个 GOP 的绝大部分而且比起其他帧 B 帧更容易被压缩而丢弃信息，这就意味着 B 帧对宽带的需求持续时间长又比其他帧更加不稳定，宽带分配还有进一步优化空间以提高宽带利用率。另外，B 帧的重要性低于其他帧，所以适当丢弃 B 帧对整个视频质量影响有限。为了验证 B 帧大小对直播视频流变化的影响，本文采用文献[6]中的方法用帧大小的变异系数 CoV(coefficient of variation)来描述视频流的速率变化。由于直播传输的视频片段时长只有几秒甚至更短，因此，本文分析 60 分钟的 H.265 视频追踪文件。对于给定量化水平参数 QP(quantizer parameter)编码后的视频帧组成的视频序列，如果此视频序列由 N 帧组成，$X_n(n=1,2,3,...,N)$ 表示视频帧的大小，那么此视频的 CoV 定义如式(1)所示。

$$CoV=\sigma/\overline{X}=\frac{\sqrt{1/(N-1)\sum_{n=1}^{N}(X_n-1/N\sum_{n=1}^{N}X_n)^2}}{1/N\sum_{n=1}^{N}X_n} \tag{1}$$

其中：σ 是标准偏差，$\overline{X}$ 是视频序列帧的中值。

本文从亚利桑那州立大学的视频追踪库[13]获取了 6 个不同的 H.265 视频描述文件用作分析，用的视频包括“Finding Neverland”“Lake House”，另外还有三个不同质量的监控视频描述文件。视频文件压缩格式为 G24B7，每个视频文件又根据 QP 不同分为三种不同质量子文件，分别是 QP=10，QP=25，QP=40。为了观察到 H.265 视频流的特征，分别计算了不同 QP 下的 BPI 帧的标准差，均值和 CoV，结果如表 I，可以观察到 B 帧的 CoV 远大于 P 帧和 I 帧的 COV，也就是说 B 帧的帧数多于 P 帧和 I 帧。根据式(1)可以知道低质量的视频 QP=40，整体上帧比较小，导致均值和标准差更小，所以 CoV 值较大，随着视频质量变高，三种类型帧的标准差，均值均变大。监控视频虽然 B 帧的 CoV 比前两个视频大，但是整体特征类似。通过均值和方差可以看出，监控视频的整体质量都要高于前两个视频，质量差异明显，便于观察差异。

对于表 2，可以看到当把视频中所有 B 帧移除，视频 CoV 明显减小，即 I 帧和 P 帧的大小变化更少。因此，降低帧大小的改变，在网络拥堵不稳定时会使 H.265 编码的视频流更加平滑。 表 2 中，没有 B 帧的视频标准差和均值差值不大，有 B 帧的视频标准差与均值相差约 3 倍，总的来说，对于画面变换较大的，比如动作电影，战争电影，标准差是决定帧大小变化的一个重要指标。对于这样的视频，除了高平均码率的要求之外，码率变化幅度也很大。相反，低质量的视频通常需要较低的平均码率。CoV 是描述码率变化的一个很好的度量指标，所以，选择性丢弃 B 帧能尽可能减少带宽需求，平滑编码视频流。

表 1　H.265 视频的 BPI 帧大小的变异系数

Table 1　CoV of B,P and I frames sizes from H.265 videos

QP	B-frames			P-frames			I-frames		
	STDEV	mean	CoV	STDEV	mean	CoV	STDEV	mean	CoV
Finding Neverland									
10	51347	101400	0.506	64148	269387	0.238	82784	337168	0.246
25	4527	3031	1.494	14196	23304	0.609	23416	43578	0.537
40	494	304	1.625	2878	3069	0.938	5170	8086	0.639
Lake House									
10	54241	88444	0.613	84243	232496	0.362	112836	291293	0.387
25	3703	2233	1.658	18177	19639	0.926	40271	46008	0.875
40	354	226	1.566	2817	2485	1.134	8268	8816	0.938
this paper Video									
10	131385	189778	0.692	181907	411579	0.442	209239	527882	0.396
25	15471	8582	1.803	37002	46454	0.797	66866	94784	0.705
40	1022	582	1.756	5483	6505	0.843	19250	20422	0.943

表 2　丢弃 B 帧视频 CoV 的变化

Table 2　The cov difference after discarding B-frames

QP	B-frames			Without-B-frames		
	STDEV	Mean	CoV	STDEV	Mean	CoV
Finding Neverland						
10	83878	125222	0.67	77770	291980	0.266
25	12202	6410	1.904	20210	30062	0.672
40	2207	858	2.572	4475	4741	0.944
Lake House						
10	81983	108900	0.753	98704	252095	0.392
25	14193	5508	2.577	30253	28429	1.064
40	2611	772	3.382	6081	4596	1.323
this paper Video						
10	165792	222355	0.746	199115	450383	0.442
25	29891	15332	1.95	54056	62579	0.864
40	6037	1903	3.172	13661	11149	1.225

3　模型建立和算法实现

根据 H.265 编码标准，GOP 中的 B 帧是参考 I 帧和 P 帧得到的，因此这些帧的大小可能存在很强的相关性。文献[6,12]用式(2)计算了 MPEG-4，H.264，H.265 编码视频 GOP 中的每个 B 帧大小与每个 I 帧或 B 帧的相关系数 ρ_{xy}，同时通过式(3)计算了 B 帧的自相关性。结果表明 B 帧之间有很强的自相关性，以及 B 帧和 P 帧之间也有很强的相关性，所以可以根据这些信息更加准确的预测 B 帧。

$$\rho_{x,y}=\frac{E(X,Y)-E(X)E(Y)}{\sigma_x\sigma_y} \tag{2}$$

其中：X 表示 B 帧大小，Y 表示 I 帧或 P 帧大小，$E(X,Y)-E(X)E(Y)$ 为 X 和 Y 的协方差，σ_x 和 σ_y 是 X 和 Y 的标准差。

$$r(k)=\frac{E[(X_n-\overline{X})(X_{n+k}-\overline{X})]}{\sigma_x^2} \tag{3}$$

其中，X 是 B 帧的大小，σ_x 是 X 的标准差，k 是滞后帧数。E(XY)-XY

由于 H.265 编码的视频流里有大量随机产生的小的 B 帧[9,12]，为了提高 B 帧预测的准确度，参考文献[6,12]提出的对 B 帧分类的方法，将 H.265 视频中每个 GOP 包含的 B 帧以所有 B 帧大小中值为分界线分为两个子集 SBF(small B-frames)和 BBF(big B-frames)，本文用寻找最佳分布拟合软件 EasyFit 对大量 H.265 编码视频的分析，发现 H.265 视频中的小帧 SBF 的最佳分布是对数逻辑分布或者皮尔逊 V 分布，大帧 BBF 的最佳拟合分布大部分是韦伯尔分布，较少服从均匀分布，与文献[12]一样，用最大似然估计方法确定这些分布的参数。在模型训练阶段，为了找到适合 SBF 和 BBF 的最佳分布，用相对百分比误差 (relative percentage error, RPE)[6]去评估 B 帧的预测值与真实值之间的误差。本文实验用的 H.265 视频 GOP 模式为 G24B7 如表 3 所示，寻找 GOP 中的 B1~B21 帧最佳分布，用来预测随后到来的 GOP_t 中的 B 帧。

表 3　G24B7 的 GOP 编码模式

Table 3　The GOP encoding pattern of G24B7

I_t B_t1 B_t2 … B_t6 B_t7 P_t1 B_t8 B_t9 … B_t13 B_t14 P_t2 B_t15 B_t16 … B_t20 B_t21

$$RPE=\frac{\sum_{m=1}^{L}\varepsilon_m}{\sum_{m=1}^{L}X_m}x100\% \tag{4}$$

其中 $\varepsilon_m=\hat{X}_m-X_m$ 为边缘误差，即预测值与真实值之间的差。

除此之外，为了更加准确的找到最佳分布，本文还用了均匀绝对百分比误差 MAPE(mean absolute percentage error)如式(6)以及 QQ-plot 进行误差评估，QQ-plot 是一种强大的拟合优度测试比较图形[14]，它可以直观得验证一组数据是否来自同一分布。

$$MAPE=\frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{A_t-F_t}{A_t}\right| \tag{5}$$

其中 A_t 为真实值，F_t 为预测值。

综上所述，提出了预测 B 帧寻找最佳分布的训练模型算法如算法 1 所示。

算法 1　B 帧寻找最佳分布训练算法.

```
//m 表示每个 GOP 里 B 帧的数量
//ArithmeticalMean 是 GOP 里所有 B 帧大小的算术平均值
//XPαb 和 XPβb 表示 X 分布关于 BBF 的参数
//XPαs 和 XPβs 表示 X 分布关于 SBF 的参数
Procedure FindingTheBestDistribution
  for CountBFrame=0 to m by increment do
    ArithmeticalMean from Bframes;
            if( Bframe < ArithmeticalMean )
      SBF←Bframe;
    else BBF←Bframe;
    end if
    //根据各种分布产生随机数
    LogRB=LogisticRegression(LPαb， LPβb);
    LogRS=LogisticRegression(LPαs， LPβs);
    PearsonVB=PearsonV(PPαb，PPβb);
    PearsonVS=PearsonV(PPαs，PPβs);
    WeibB=Weibull(WPαb， WPβb);
    WeibS = Weibull (PPαs，PPβs);
    UniformB=Uniform(UPαb，UPβb);
    UniformS= Uniform (UPαs，UPβs);
    …//其他分布
    // DistrbTypeNum 表示用了多少种分布
    //CriterionListX 表示存储各种评估值的链表
    // DistrbTypeX 表示以上分布产生关于 xBF 的随机数
    for i= 0 to DistrbTypeNum by increment do
      CriterionListB[RPEType][CountBFrame]←RPE(DistrbTypeB[i]，BBF);
      CriterionListS[RPEType][CountBFrame]←RPE(DistrbTypeS[i]，SBF);
      CriterionListB[MAPEType][CountBFrame]←MAPE(DistrbTypeB[i]，BBF);
      CriterionListS[MAPEType][CountBFrame]←MAPE(DistrbTypeS[i]，SBF);
      CriterionListB[QQplotType][CountBFrame]←QQplot(DistrbTypeB[i]，BBF);
      CriterionListS[QQplotType][CountBFrame]←QQplot(DistrbTypeS[i]，SBF);
    end for
  end for
  //Choose the best distribution
  for CountBFrame =0 to m by increment do
    BestDistBBF(CountBFrame)←Best(CriterionListB);
    BestDistSBF(CountBFrame)←Best(CriterionListS);
  end for
            Get(Pbb，Pbs，Psb，P ss);
end Procedure
```

通过模型训练为 BBF 和 SBF 找到最佳分布以后，在此后预测下一个或更多 B 帧的大小之前，需要判断即将到来的 B 帧是 BBF 还是 SBF，需要以上训练结果进行对 H.265 编码的视频 B 帧类型（BBF 或者 SBF）实时预测，最后根据相应的分布得到 B 帧大小的预测值。本文提出了一种基于马尔可夫模型的预测算法，为 GOP 的每一个 B 帧建立一条 2 状态马尔可夫链，比如本文用到的 GOP 模式为 G24B7，那么一个 GOP 里包含 21 个 B 帧，就需要 21 条不同 2 状态马尔可夫链。 显然，每个状态链有 4 个转移概率:BBF→BBF，BBF→SBF，SBF→BBF，SBF →SBF，用式(6)描述。接着定义 4 个转移概率分别用 P_{bb}，P_{bs}，P_{sb}，P_{ss} 来表示由算法一在训练时得到，图 1 表示第 t 个 GOP 中的第 m 个 B 帧通过转移概率预测第 t+1 个 GOP 中的第 m 个 B 帧的类型。以预测 B1 帧为例，分为两步：第一，计算出需要预测 B1 帧的类型，第二，用最佳分布产生 B1 帧大小，如图 2 所示，由此提出算法二。

$$P_{ij}(t,t+n)=P\{X_{t+n}=a_j\,|\,X_t=a_j\}\quad t=1,2,\dots,N \tag{6}$$

其中: t 表示第 t 个 GOP， a_i 表示第 t 个 GOP 时的 B 帧类型，a_j 表示第 $t+n$ 个 GOP 时的 B 帧类型，P_{ij} 表示在第 t 个 GOP 时状态 a_i，在第 $t+n$ 个 GOP 转移到状态 a_j 的概率。

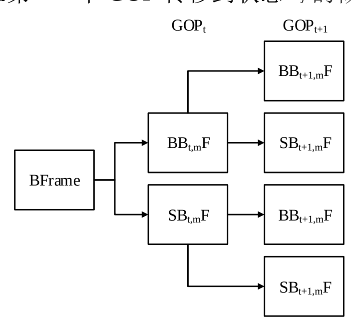


图 1　马尔可夫链预测 B 帧类型

Fig. 1　Use Markov chains to predict the type of B frame

算法 2　一种基于马尔可夫琏的 B1 帧实时预测算法。

```
// GOPn 表示 GOP 的数量
//(Pbb，Pbs，Psb，P ss )为转移矩阵，由训练算法得到
Procedure B-FramesPrediction
  for t=InitalNum to GOPn by increment do
     //计算 B 帧为 BBF 和 SBF 的概率
     (Pt+1,b，  Pt+1,s) ← ( (Ptb，  Pts)(Pbb，Pbs，Psb，P ss ) )
     BestDistribution of BBF or SBF←Max(Pt+k,b，  Pt+k,s)
     //通过最佳分布预测 B 帧大小
     Bframe ← BestDistribution()
   end for
end Procedure
```

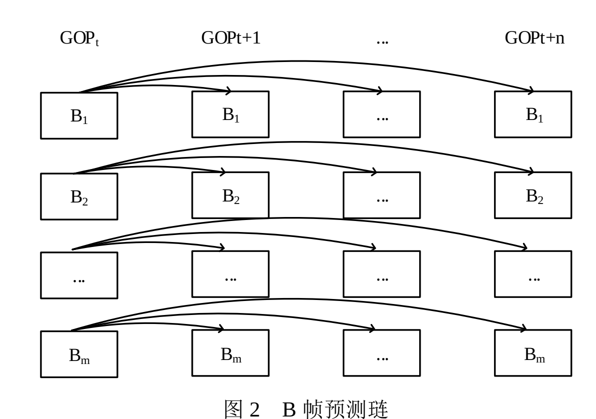


图 2　B 帧预测琏

Fig. 2　Markov prediction chains for B frames

4　模型优化

定义 1　B 帧预测值 *bp* 与真实值 *ba* 之间的误差，记为 E。

$$E=\sum_{k=1}^{n}\left|bp_k-ba_k\right|^2 \tag{7}$$

算法 2 通过初始模型的训练获得 B 帧两种类型 BBF 和 SBF 的转移矩阵，在 GOP 序列里预测 $B_{t,m}$，$B_{t+1,m}$，$B_{t+2,m}$，…，$B_{t+n,m}$ 帧(t+n 表示 GOP 序列号，m 表示一个 GOP 里 B 帧序列号)用的转移矩阵与初始矩阵一致，而视频直播过程中，随着时间推移可能不满足马尔可夫琏转移概率的时齐性，为了尽可能避免这种情况，本文提出了由阈值 FT 来决定是否重新训练，因此，FT 值的大小将决定训练频率，及当 FT 值较大时，预测过程中训练次数将增多，算法消耗资源较多。反之，当 FT 值较小时，预测过程中训练次数将减小。在此模型中，当 $E>FT$ 时开始再次训练算法一，其中 E 由式(6)得到。又因为，每次训练后原来的 FT 可能不适用当前预测模型，所以需要 FT 与 B 帧预测的准确度建立联系，既不让 FT 过大过小，也不让预测准确度波动太大。

引入步长调节因子 AF 及在预测过程中限制向前预测 B_m 帧帧数，将有效解决这一问题，令 $FT=1/ExAxFunc(AF)$，E 为积累误差，A 为系数，$0<AF<1$，$0<Func(AF)<1$，Func(AF) 使得当 AF 值较小时，FT 取值较大，及预测算法执行次数增加，积累误差肯定会比执行算法次数少时大，所以应该加大再次训练容忍度即 FT 变大。反之，FT 取值较小。因此，一种可能取值 $Func(AF)=(1-AF)^2$ 及 $FT=1/ExAx(1-AF)^2$。

为了根据网络带宽选择性丢弃 B 帧，参考 TCP 阻塞控制算法 Westwood[15]的带宽估计方法，每个 RTT(Round-Trip Time)时间内，对发送端做带宽估计，即 RTT 内成功被 ACK 了多少字节。最后用带宽估计值来决定 B 帧的丢弃与否，如式(8)所示，d_k 表示是否丢弃 B 帧的结果。

$$d_k=\begin{cases}丢弃 & \eta_k<bp\\ 保留 & \eta_k>bp\end{cases} \tag{8}$$

综上，改进后的测试模型用算法三来描述，

算法 3　一种新的 B1 帧实时预测算法。

```
// GOPn 表示 GOP 的数量
// Pt+k, b，  Pt+k, s 为第 t+k 个 GOP 大小帧产生的概率
//(Pbb，Pbs，Psb，P ss )为转移矩阵，由训练算法得到
//PredictN 表示在当前 GOP，需要向前预测 B1 帧帧数，及预测步
长，由步长调节因子 AF 的到
Procedure B-FramesPrediction
  for t=InitalNum to GOPn by increment do
     //计算 B 帧为 BBF 和 SBF 的概率
     for k=1 to PredictN by increment do
        (Pt+k, b，  Pt+k, s) ← ( (Pt, b，  Pt, s)(Pbb，Pbs，Psb，P ss ) )
        BestDistribution of BBF or SBF←Max(Pt+k,b，  Pt+k,s)
        //通过最佳分布预测 B 帧大小
        Bframe ← BestDistribution()
        if (dt=1) drop Bframe
        else E = E+|Bframe-RealValue|^2
        end if
     end for
     t=t+ k
  end for
  E=mean (RPE，MAPE，QQplot)
  if (E > FT)
    FT= 1/E x A x (1-AF)^2.
    FindingTheBestDistribution()
  end if
end procedure
```

5　实验与仿真

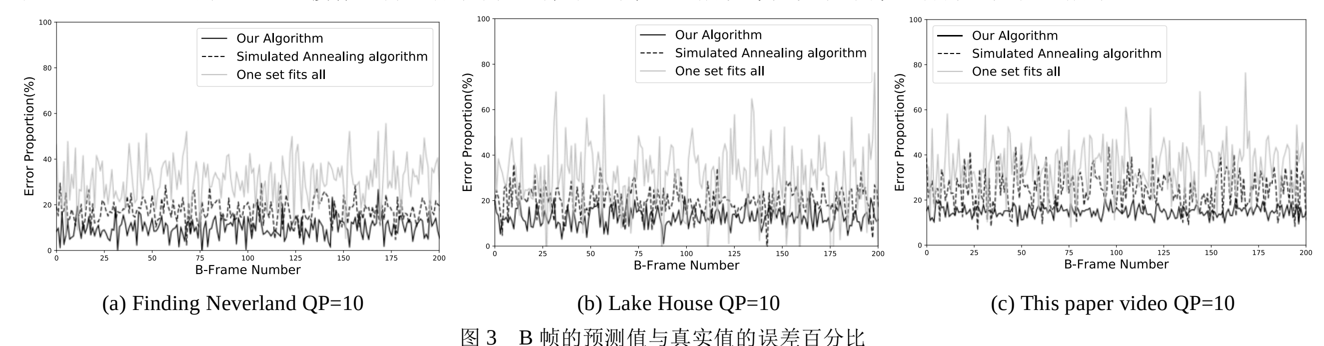
5.1　算法仿真

为了评估本文提出 VBRP 算法性能，开发了 B 帧预测程序，对"Finding Neverland""Lake House"和监控视频中 1000 个 B 帧进行预测。 质量为 QP=10 的视频文件作为预测程序的输入，输出为预测的 B 帧序列，来模拟直播视频流。本实验验证本文提出的算法实际预测误差大小百分比。选择 Intel Core i5-3470 CPU ，16GB DDR3 Centos6.8 操作系统的服务器作为物理机实验平台。如图 2 所示，x 轴为 B 帧的序列号，y 轴为预测值与真实值的绝对误差百分比，在积累误差较大时，本文提出的 VBRP 算法比模拟退火预测算法和 One Set fits all 更加平稳，在预测"Finding Neverland"，"Lake House"和监控视频中的 B 帧时，VBRP 算法(图 3 中黑色曲线)误差率在 3%~25%之间，显著优于其他两种算法。

5.2　系统实现

目前直播系统解决方案很多，其中比较出色的有 Adobe 公司的 RTMP 协议直播方案，Apple 公司的 HLS(HTTP live streaming)协议直播方案和谷歌的 WebRtc 方案。RTMP 方案比较成熟，延迟低，搭建方便，但是需要通用性不够。HLS 方案适合各种浏览器直接播放，兼容性强，特别是移动终端观看视频直播。WebRtc 方案延迟低，兼容性基于前两者之间，实现较复杂，并发性低。因此，为了实验的便利性，本文参考文献[16~18]实现了基于 RTMP 协议的视频直播系统。由三部分组成：Centos 作为服务器的基本操作系统，Nginx + Nginx-rtmp-module 作为流媒体发布服务器，FFmpeg 作为推流服务器。

如图 3 所示，在 Centos6.8 系统上，编译安装 Nginx 和 Nginx-rtmp-module，修改 Nginx 的配置文件，使其支持 RTMP 协议的流媒体发布。接着，编译安装 FFmpeg 并在编译时开启 H.265 协议的支持，如 –enable-libx265。 最后开始用 FFmpeg 进行推流到流媒体服务器的缓存路径，客户端通过 ip 地址来观看直播视频。然后对预测算法运行的拓扑位置进行分析，第一种方法，预测算法整合进 Nginx 流媒体服务器中，重组 RTMP 的 Chunks 缓存，再进行预测，对随后到来的 Chunks 执行：重组→丢弃 B 帧→拆分。 第二种方法，将预测算法整合进 FFmpeg，训练模型单独创建一个训练线程，在生成 Chunks 前根据算法执行 B 帧丢弃：丢弃 B 帧→拆分。第一种方法增加了延时，使得预测滞后，影响了整个预测系统。第二种方法，在编码 H.265 之后生成 RTMP 的 Chunks 之前进行预测，然后根据网络带宽执行 B 帧丢弃。显然，相对于第一种方法，第二种方法延迟更小，消耗的资源更少。所以本文实现了第二种方法如图 4 所示。



(a) Finding Neverland QP=10　　(b) Lake House QP=10　　(c) This paper video QP=10

图 3　B 帧的预测值与真实值的误差百分比

Fig. 3　Percentage error between the predicted value and the real value of the B frames

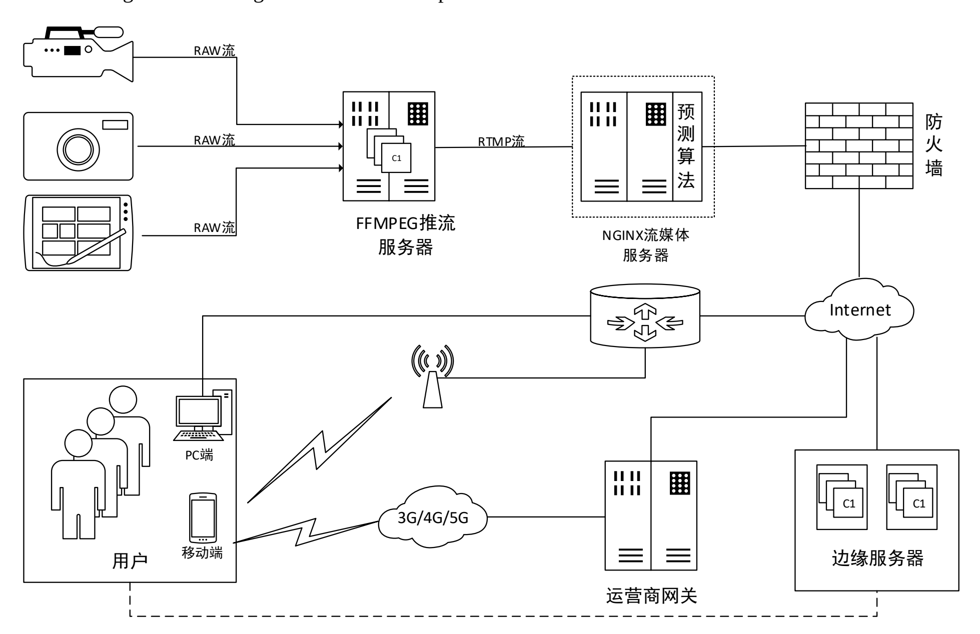


图 4　预测算法运行系统

Fig. 4　System of prediction algorithm running

6　结束语

本文提出的 VBRP 模型通过研究 H.265 编码的视频流特征，根据特征将 B 帧分为 BBF 和 SBF 两种类型，在训练阶段寻找到适合 BBF 或 SBF 帧类型的最佳拟合分布。用一种新的马尔可夫琏预测即将到来的 B 帧类型，确定类型后，最佳拟合分布产生的 B 帧就是预测帧大小。同时引入了 FT 作为模型训练的阈值，AF 作为 B 帧预测的步长调节因子，使模型训练频率动态根据 AF 和积累误差 E 动态变化，进一步提高了预测模型的准确率。在模型仿真验证之后，将 VBRP 算法应用于 Nginx 和 FFmpeg 等组成的流媒体发布系统。实验表明，所提出的 VBRP 算法要优于模拟退火预测算法和静态视频文件预测算法[12]。然而，提出的模型 AF 的值是经验性的，有待进一步改进，选择性丢弃 B 帧所依据的网络带宽预测仅仅是根据 RTT 来确定的，由 RTT 的变化在一定程度上反映了网络拥塞程度的变化，B 帧的丢弃不能完全依赖 RTT 获取的带宽值，否则 B 帧的丢弃将失去合理性。下一步工作，将研究准确度更高的网络带宽预测模型并将应用到模型中，并分析网络带宽大小对 AF 取值的影响，使 AF 的值随网络带宽大小变化，进一步提高预测准确度。

参考文献：

[1] Annual Meeting of Stockholders of Netflix, Inc. [EB/OL].(2017) [2018-07-10]. https://s22.q4cdn.com/959853165/files/doc_financials/annual_reports/0001065280-18-000069. pdf.

[2] Cisco visual networking index: global mobile data traffic forecast update, 2015–2020 [EB/OL]. (2016)[2018-07-10]. https://www. cisco. com/c/dam/m/en_in/innovation/enterprise/assets/mobile-white-paper-c11-520862. pdf

[3] J Sullivan G, Ohm J R, Han W J, *et al.* Overview of the high efficiency video coding (HEVC) standard [J]. IEEE Trans on Circuits and Systems for Video Technology, 2012, 22 (12): 1649-68.

[4] Qi Xin, Yang Qing, Nguyen D T, *et al.* A context-aware framework for reducing bandwidth usage of mobile video chats [J]. IEEE Trans on

Multimedia, 2016, 18 (8): 1640-1649.

[5] Sarkar U K, Ramakrishnan S, Sarkar D. Modeling full-length video using Markov-modulated gamma-based framework [J]. IEEE/ACM Trans on Networking, 2003, 11(4): 638-649

[6] Lanfranchi L I, Bing B K. MPEG-4 bandwidth prediction for broadband cable networks [J]. IEEE Trans on Broadcasting, 2008, 54(4): 741-751.

[7] Parmar H, Thornburgh M. Real-Time Messaging Protocol (RTMP) specification (version 1. 0) [EB/OL]. (2012)[2018-07-10]. http://wwwimages.adobe.com/www.adobe.com/content/dam/acom/en/devnet/rtmp/pdf/rtmp_specification_1. 0. pdf.

[8] Miller K, Bethanabhotla D, Caire G, *et al.* A Control-theoretic approach to adaptive video streaming in dense wireless networks [J]. IEEE Trans on Multimedia, 2015, 17 (8): 1309-1322.

[9] Kalampogia A, Koutsakis P. Using simulated annealing for improved video bandwidth prediction [C]//Proc of IEEE Conference on Computer Communications Workshops.2017: 701-705.

[10] Colonnese S, Russo S, Cuomo F, *et al*. Timely delivery versus bandwidth allocation for DASH-based video streaming over LTE [J]. IEEE Communications Letters, 2016, 20(3): 586-589.

[11] Sodagar I. The MPEG-DASH standard for multimedia streaming over the Internet [J]. IEEE Multimedia, 2011, 18(4): 62-67.

[12] Kalampogia A, Koutsakis P. H.264 and H.265 video bandwidth prediction [J]. IEEE Trans on Multimedia, 2017, PP (99): 1-1.

[13] Arizona State University. Trace files and statistics [EB/OL]. (2012) [2018-07-10]. http://trace. eas. asu. edu/tracemain. html.

[14] Ito M, Yoshida K, Hachiya H, *et al.* Quantification of the scatter distributions for liver fibrosis using modified Q-Q probability plot [C]// Proc of Ultrasonics Symposium. 2014: 2394-2397.

[15] Mascolo S, Casetti C, Gerla M, *et al.* TCP westwood: end-to-end bandwidth estimation for efficient transport over wired and wireless networks [C]//Proc of ACM MOBICOM.2001.

[16] Zhao Pengyu, Li Jianwei, Xi Jianxiao, *et al.* A mobile real-time video system using RTMP [C]//.Proc of the 4th International Conference on Computational Intelligence and Communication Networks. 2012: 61-64.

[17] Huang Jian, Wu Dongmei, Liu Xiaopei. Implementation of the Rtmp Server Based on Embedded System [C]//Proc of International Conference on Computer Science and Information Processing. 2012: 160-162.

[18] Lei Xiaohua, Jiang Xiuhua, Wang Caihong. Design and implementation of streaming media processing software based on RTMP [C]// Proc of International Congress on Image and Signal Processing. 2013: 192-196.